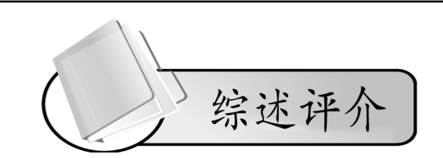


自动引文摘要研究述评

刘天祎　步　一　赵丹群　黄文彬
(北京大学信息管理系　北京　100871)

摘要: 【目的】对引文摘要领域的国外主流研究方法和步骤进行综述分析。【文献范围】选取 2007 年以来引文摘要领域的重要研究及此前自动摘要、引文分析领域的研究进展。【方法】基于文献调研，介绍该领域的基本概念以及自然语言处理的方法在引文摘要中的应用。【结果】引文句在摘要实践中起到重要的概括作用、指示作用和关联作用，具有一定的优越性。【局限】缺乏对引文摘要领域现有成果和可能达成的理想情况的比较。【结论】引文摘要拓展了自动摘要和传统的信息计量学的研究方向，并对改进自动摘要原有的评估方案提出要求，同时产生了有关引文窗口扩展、语料库构建等一系列新问题。本文对这些问题进行探讨，并对引文摘要未来的研究发展进行展望。

关键词: 自动摘要　引文摘要　引文句　自然语言处理
分类号: G350

1　引　言

自动摘要是自然语言处理的重要研究课题之一，长期以来，其研究主要关注基于正文文本信息的摘要生成技术和方法。完全基于正文的摘要(包括作者摘要)虽然能够较好地反映原文内容，但在对原文影响力的概括能力上则比较有限，更无法反映文献影响力的历时性变化[1]。

引用(被引用)关系是学术文献之间最有研究价值的关联关系之一，可看作是从其他学者视角对被引文献内容的一种概括或解读，进而体现被引文献对其他研究的学术影响或价值。因此，借鉴文献计量学中的引文分析思想，开展基于引文信息的摘要研究(简称“引文摘要”)逐渐成为自动摘要领域最近 10 年来一个新的探索方向。

目前，引文摘要研究的基本思路可以概括为：搜寻目标文献的所有施引文献的全文，以获取全文中对目标文献引用标注所在的所有句子或其他相关信息，将其看作一个集合，然后从这个集合中选取一个子集加工、生成目标文献的摘要，并保证这个子集具有足够的压缩率和较好的概括能力。为此，它的研究步骤(或关键问题)主要包括：选择合适的、可获取全文的语料库；引文句(域)识别与摘取；引文类型和引文目的识别，对引文句进行分类和筛选；引文句的组织和排序，形成引文摘要(初稿)；摘要后处理；摘要评估。

自动摘要原本就被广泛应用于信息检索。Bradshaw 认为引用关系可以用于学术检索效果的改进，这反映了引文摘要的优越性[2]。此外，研究表明对于高被引文献来说，引文摘要更具客观性和多样性[3]，并在揭示目标文献信息方面具有明显的优越性[4-5]。此外，与基于正文的摘要方法相比，引文摘要不仅在内容上比原文句子更具概括性，由于经过了一轮人工的分析，还具有一定的评论性和延伸性，能够更好地反映出原文中有重要意义的部分。

事实上，引文摘要研究的兴起，主要得益于全文语料的日益普及和自然语言理解技术的进步，同时，也可将其视为引文语境分析(Citation Context Analysis)技术的一类重要应用。自 2008 年 Qazvinian 等首次开

通讯作者：赵丹群，ORCID：0000-0003-0685-6689，E-mail：zdq@pku.edu.cn。

展引文摘要试验研究以来[6]，这一“自动摘要”与“引文分析”相互交叉融合发展起来的新研究方向不仅得到了学者的广泛关注，而且取得了显著的研究进展，并已从初期过于偏重对文献计量学及引文分析方法的借鉴应用，逐步转向于更多关注自然语言处理技术、特别是文本语义与情感分析等新技术的应用。不过，受限于低被引文献引文语料信息的匮乏，以及结构化全文语料广泛获取的相对困难性(医学领域除外)，对引文摘要技术方法更深入的研究、推广及应用还存在较大的拓展空间。

在 Web of Knowledge 平台上以检索式 TI = (citation* OR reference* OR bibliography) AND TI = (summar* OR survey OR extract* OR abstract)进行检索，并针对研究方向进行精炼，得到 2007 年–2015 年的文献 300 余篇，数量较少，其中与引文摘要直接相关的更仅有数十篇。可以看出针对引文摘要的学术研究尚处于起步阶段。国内针对自动摘要领域的研究比较有限[7]，根据对 CNKI、万方等数据库的检索结果来看，国内对引文摘要领域的研究几乎是一片空白。本文通过较全面的文献调研，尝试从(单文档)引文摘要研究的关键步骤入手，对这一新兴研究课题的国外研究进展进行较为系统的分析评述，并针对一些基本概念加以厘清，以期为国内引文摘要研究的深入发展提供必要的借鉴和启迪。

2 引文域识别和引文窗口扩充

引文摘要研究主要建立在对文献引用与被引用关系中蕴含的主题相关性及价值进行挖掘利用的认知基础上，因此引文域识别与提取不仅是其研究中的关键步骤之一，也是首先面临并需妥善解决的一个问题，它对后续步骤的实施及摘要生成具有重要的支撑作用。

早在 2004 年，Nakov 等就提出“Citance”(即 Citation Sentence，引文句)这一新术语[8]，意指施引文献中围绕在引用标记(符号)周围的句子。狭义的“引文句”可理解为引用标记所在的句子本身，广义理解可扩展为引用标记所在句子及其周围信息，也即引文上下文(Citation Context)或引文语境信息。通常情况下，引文域可以看成施引文献中论及被引文献的语段，并且这种论述应该是比较明确的、有意义的，同时具有可以识别的边界或阈值。引文域的范围，可称之为“引文窗口”(Citation Window)。

在早期的研究中，引文域的识别和抽取多基于狭义的理解，一般通过对特征边界的识别直接对引用标记所在的句子本身进行抽取。例如，Nanba 等通过人工方式，先期得到了与引文域识别相关的 86 个线索词，并据此完成引文域的提取任务[9-10]。近年来更常见的方法是使用正则表达式来描述各种引用方式的模式。最近，研究人员则开始采用扩充引文窗口的方法提取相关信息，即尝试利用引文句附近，或者其他语义上相接近的句子，以有效提高用于引文摘要信息的丰富度。

引文窗口的扩充对了解论文引证中的引文习惯、引文原因以及文献信息流向特点等具有重要意义[11]。目前，引文窗口扩充(或扩展)的常用方法主要有两种：基于距离扩展和基于相似关系扩展。其中，基于距离的扩展方法较为简单，一般可通过计算其他句子与引文句的物理距离来给它们赋予不同的权重，从而考虑是否将其纳入到引文域中[1]；或者是，直接指定引文窗口中可包含的具体句子(或段落)的数目。而基于相似关系的扩展方法则相对复杂，它致力于将与引文句语义相关性较高的语句信息(在位置上并不一定相邻)识别出来并纳入到引文域中，这种方法尤其适用于“隐式引用”(Implicit Citation)情形[12]。例如，当学术文献中对所引用文献采用哈佛体系进行著录或标记时，往往就会产生或带来较多的隐式引用。对此，Athar 等的研究认为，隐式引用往往包含更加丰富的语义信息，并有较大的利用价值[13]，有必要将其扩充到引文窗口中。

引文域的摘取质量直接决定引文摘要的质量。目前，引文域的识别在技术上主要建立在对引用的边界特征的提取上；而引文窗口的扩充，则更多依赖于对自然语言文本的语词特征和语义特征的分析理解技术，未来还存在较大的改进余地或优化空间，尤其是针对引文域中句子的主题、观点识别等问题，尽管目前还少有研究将其与引文摘要相结合，但也是非常有意义的改进方向。

3 引文句分类

引文句分类旨在对从所有引文域中摘取的某目标文献的引文句(集合)按照一定的结构(或逻辑)标准进行组织整理。例如，按照研究目的、研究方法、研究缺陷、研究结论等对引文句进行分组或分类，并据此

对引文句进行初步的筛选和剔除，保证摘要的全面性和简洁性，同时也便于按照其内在的逻辑顺序加以组织，以保证摘要的可读性。

早期，引文句分类主要沿用对作者引用行为或引用动机方面的一些研究成果及研究思路，包括作者对被引文献的评价是积极或消极等，能够在一定程度上反映施引文献和被引文献之间的关系。例如，Nanba 等基于 1965 年 Garfield 对作者引用行为总结出来的 15 种类型，将引文句概括分为 TypeB/C/O 三种类型[10]。这类研究中，早期的研究多依托特征词、线索词的识别，最近的研究更多地引入情感分析技术。

新兴的文摘结构理论研究为引文句分类提供了新的可能。这类理论最早是通过对学术文献摘要句子所属的文章区块进行研究和分类。具体的研究工作主要包括：研究文档本身或文档摘要的结构，为自动摘要提供指导；研究在某种文摘结构下的分类技术与分类特征。其中，前项研究往往是后项研究的理论基础，以便确保分类结果能够使引文句出现在摘要的合适位置，从而符合读者的阅读习惯(逻辑)。

基于文章自然分段进行分类是非常直观的，也在许多研究当中得以应用，白光祖等利用朴素贝叶斯算法，在小样本量的条件下按照学术文章的分段名称进行识别，取得了比较好的效果[14]。然而不同文献的自然分段往往不同，这导致了该方案的天然缺陷。由此，一些新的分类理论得到了发展。Teufel 等提出了非常严密的论述结构(Argument Zones，AZ)模型[15-16]。AZ 理论最初将文章中句子的修辞地位(Rhetorical Status)分成 7 类：目标、结构、作者自己的观点、背景、对比、基础和其他，并在后续的修正版本 AZ-II 中进一步细化成 15 个部分。这种细化处理被认为更加富含信息量，并且对不同学科的适应能力更强[16]。此外，还有一种比较重要的理论——核心科学概念(Core Scientific Concept，CoreSC)结构理论[17]，其分类比 AZ 理论还要更加细致。对上述三种分类理论(自然分段、AZ 理论、CoreSC)的比较研究发现，利用机器学习对它们提供的类目进行摘要句分类的效果都比较可靠，并且彼此的类目之间还具有潜在的联系[18]。近年来一些比较重要的学术文摘结构分类研究如表 1 所示：

表 1　学术文摘结构研究的相关理论

第一作者/年份	理论名称	模型	特征	分类方式
Mizuta/2006[19]	Zone Classes	第一组：背景、问题、作者自己的观点(方法、结论、见解、暗示、意义、其他) 第二组：关联、区别 第三组：大纲	词语、主要动词、时态、情态、分块或段落、句子、引用、划线部分、句子宾语	决策树
Teufel/ 2006[20]	Argument Zoning (AZ)	目标、结构、作者自己的观点、背景、对比、基础、其他	分块结构、段落结构、标题、句子长度、词频*反文档频率、内容、动词时态、动词情态、引用等	朴素贝叶斯分类器
Ehrler/2005[21]		意图、方法、结果、结论	距离、词频*反文档频率(tf*idf)	向量空间分类器、正则表达式匹配分类器
Hiroakata/2008[22]		目标、方法、结果、结论、两个自定义的前缀类	N 元语、相关句子、句子位置、引文上下文特征	条件随机场、支持向量机
Teufel/2009[23]	Argument Zoning-II (AZ-II)	在原版基础上加以细化	人工标引员	人工标引员
Liakata/2010[24]，Liakata/2013[17]	Core Scientific concepts (CoreSC)	假设、动机、背景、目标等 18 种类别	人工标引员	人工标引员

Contractor 等将 AZ 理论作为引文句分类和筛选两个步骤的特征，取得了比较好的效果[25]。总体而言，针对文摘结构理论的研究正在逐渐避免主观性，而更多地依赖于分类特征的选择。目前的研究更加重视文摘结构理论在不同学科的普适性，而这也依赖于更加具有普适性的特征选取和处理技术。

4　引文句组织和排序

引文句组织和排序是指从经过整理的候选引文句集合中筛选出内容表达能力最强的句子，并按照一定

方法对其进行排序，从而生成概括性强、逻辑连贯的引文摘要。

目前，引文句组织和排序的主要研究工作包括：

(1) 利用一定的相似性度量方式，将引文句进行聚类。聚类一般是基于引文句之间的内容相似性进行处理，以剔除表意相似的句子。常用的聚类方法有：层次聚类[4,26]；最大边缘相似度(Maximum Marginal Relevance, MMR)及其变体，一种常用的基准算法[27-30]；MEAD[31]及其变体[1, 4, 26, 28, 32-33]。其中，MEAD是一个开放的免费自动摘要研究平台，提供有多种基本的摘要算法系统，并经常在评估过程中作为基准被研究人员使用，其变体包括MEAD-Centroid、MEAD-LexRank[34]。另外，该系统在多文档摘要中也有非常广泛的应用。

此外，由于引文句和原文之间、引文句和引文句之间自带引用关系，能够形成引用网络，因此引文句聚类又格外适用图模型摘要问题(IGS Problem), Shi等对大规模的引文网络图进行压缩，实现了对研究主题变迁的追踪和重要性揭示[35]。尽管该研究不是一个典型的引文摘要问题，但是其适用的图模型思想、可视化方法和算法改进却对引文摘要领域有借鉴意义。

(2) 对引文句的重要性进行打分(加权)，以便排序输出。例如，Mei等通过设计一种基于影响力的打分方案，对能反映文档影响的引文句进行排序[1]; Qazvinian等则基于对引文句中的关键词提取及关键词的出现密度和重要性等信息，对引文句进行排序，并据此实现引文句的去重处理[28]。目前完全依赖于对引文句进行打分的研究正在减少，由于打分一般依赖于句子中语词，例如特定的命名实体或者事实的出现特征，那么难以避免地，无论采取何种方案，打分高的句子往往指向的对象或者表达的意思相近，客观上影响了摘要句子的多样性。

5 摘要后处理和评估

5.1 摘要后处理

摘要后处理是指在得到摘要初稿的基础上，对已选定的引文句进行检查，包括是否存在冗余，一些重点词语的使用上是否存在指代不明、不连贯等问题。近年来，研究人员对摘要结果的后处理问题愈加关注。

目前，摘要后处理致力于解决两方面的问题：

(1) 去重。引文句表意重复的一种简单情形是语义重复，即句子对中使用的词语基本相同，意思表达也基本一致。此时去重操作比较简单，通过基于词级别的相似度计算(比较)就可以解决。而且一般而言，如果有相应的引文分类手段的话，往往也不会存在句子相似度极高的状况。而对于比较复杂的引文句表意重复问题，去重的难度就比较大。例如，两个引文句都引用了一篇有关“信息检索”的文章，但引文句一使用的词是“search engine result”，而引文句二使用“information retrieval”，其用词完全不同，但却表达了接近的语义，往往不能通过简单的相似度比较来发现，通常需要相关知识库的配合使用才能解决。

(2) 连贯性[36-37]。一般而言，引文句在格式上与正文文本是有所区别的。另外，引文句通常来自于不同的施引文献，表达习惯上也往往存在较大的差异，因此，相比基于正文的自动摘要，引文摘要方法得到的文摘的可读性和连贯性都会比较差。对此，采用的主要解决办法是：通过分析引文摘要结构和人工摘要的用词特点，插入、替换适当的代词、连词等，以利于摘要连贯性的提升[26]。

5.2 摘要评估

摘要评估主要是指针对摘要的概括性、连贯性、准确性、语法正确性、可读性等方面的定性与定量评判。传统的自动摘要评估一般采用Recall-Precision和F-measure方法，用以度量原文中被指定的模板单元被摘要所覆盖的全面性和准确性。另一个得到广泛应用的评估方案是ROUGE①，它主要基于n元语(n-gram)模型的召回率计算，其变体还有ROUGE-P、ROUGE-S、ROUGE-L等[38]。

截止目前，几乎没有专门针对引文摘要的评估方案。因此，ROUGE和Precision-Recall仍然得到了广泛使用。然而，这两种方案并不能完全适应引文摘要的评估要求。引文的用词有可能与原文有一定的偏差，难以判断其覆盖能力和准确性；由于来自于多个文档中的引文句之间彼此独立，相对于一般的自动摘要而言，引文摘要的引文句之间可能会存在更多的语义冗

①Recall-Oriented Understudy for Gisting Evaluation.

余，这种情况下，摘要可读性的降低难以避免，但现有评估方案却无法测量出来。

很多引文摘要的评估工作都基于人工参与方式进行。例如，使用与人工生成的高质量摘要进行相似度比较来评估，或者邀请领域内专家对生成的引文摘要打分。与之相近的还有问答方法[17]，即由专业人士阅读摘要内容，然后回答有关原文的问题来评测摘要内容是否抓住了原文的重点。这些方案在人力上耗费都较大，同时也难以避免人工处理的主观性。在一项近期研究中，Christensen 等使用纯人工评估方法，通过邀请专业的评估人员直接对其自动生成的引文摘要与基准摘要进行"盲评"[36]，对摘要连贯性的评估进行重点探讨。

此外，引文摘要评估还见到使用金字塔评分(Pyramid Score)的方法[39]。这里，金字塔评分是指对引文摘要标记出的"事实"(Fact)进行评分，并对标记出较多事实的摘要赋一个比较高的分数。这种评估方案将评估对象降低到语词层级，更加强调文摘中对"事实"或者"命名实体"的提取，具有一定的参考价值，但仍然存在无法评价文摘可读性、连贯性和简洁性的问题。

6 结 语

引文摘要问题的提出及相关研究工作的开展至今不足 10 年的历史。本文基于内容要素，从引文摘要的关键步骤入手，对这一新兴领域的国外研究进展进行了较为全面的文献调研和论述分析。从早期对引文句功能、作用的定性讨论，到对引文句内容、观点及倾向性的深入分析，再到基于引文句(集合)进行自动摘要的生成及评估；从单文档摘要到多文档摘要，再到自动生成基于更大规模文档的文献综述，引文摘要领域的研究工作正在渐趋成熟和深化。

然而，作为一个新兴的研究课题，引文摘要研究还存在许多不足，同时也面临着不少困难，具体表现如下：

(1) 适用的全文语料库的缺乏。目前，除 PubMed 外，其他学科基于 XML 标记的结构化全文语料还比较少见，而由于 OCR 技术的识别能力还不完美，据此得到的全文文本的噪声还比较大[30]，因此，开展引文摘要研究在很多情况下还要依赖人工方式对全文数据进行预处理，这大大影响了研究工作的效率。

(2) 缺乏有针对性的评估方案。传统的自动摘要评估方案不能直接反映引文摘要的质量，同时一些最新的有关连贯性和可读性的评估研究缺乏得到广泛认可的理论支撑。此外，评估方案过度依赖人为打分，也导致评估工作无法大规模展开。

(3) 受引文句数量、文献所属学科等因素的影响，引文摘要的长度(篇幅)差异性很大，也缺乏伸缩性。还少有研究对引文摘要合适的长度进行讨论。

(4) 引文摘要方法无法适用于在学术文献中占多数的低被引文献。

需要说明的是，本文的综述分析重点围绕单文档引文摘要而展开。近年来，在单文档引文摘要研究的基础上，多文档引文摘要以及文献综述的自动生成也得到了学者的积极关注。由于后两种摘要的生成难度明显高于单文档摘要，特别是在引文句分类、引文句聚类、引文句语义去重及排序等方面，技术难度及挑战性更高[9,10]。因此，在未来的研究中，如何对引文句这类语料进行更为深入的语义理解和情感分析，如何基于引文句的指示与关联作用在单、多文档摘要之间形成联动关系，以及如何与基于正文的自动摘要技术相互融合，都将成为引文摘要研究下一步的努力方向。

参考文献：

[1] Mei Q, Zhai C. Generating Impact-Based Summaries for Scientific Literature [C]. In: Proceedings of ACL-08: HLT, 2008: 816-824.

[2] Bradshaw S. Reference Directed Indexing: Redeeming Relevance for Subject Search in Citation Indexes [C]. In: Proceedings of the 7th European Conference on Research and Advanced Technology on Digital Libraries (ECDL 2003), Trondheim, Norway. Springer, 2003: 499-510.

[3] Elkiss A, Shen S, Fader A, et al. Blind Men and Elephants: What do Citation Summaries Tell Us about a Research Article? [J]. Journal of the American Society for Information Science and Technology, 2008, 59(1): 51-62.

[4] Mohammad S, Dorr B, Egan M, et al. Using Citations to Generate Surveys of Scientific Paradigms[C]. In: Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2009: 584-592.

[5] Kan M-Y, Klavans J L, McKeown K R. Using the Annotated Bibliography as a Resource for Indicative Summarization [C].

In: Proceedings of LREC, Las Palmas, Spain. 2002: 1746-1752.

[6] Qazvinian V, Radev D R. Scientific Paper Summarization Using Citation Summary Networks[C]. In: Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1, 2008: 689-696.

[7] 王连喜. 自动摘要研究中的若干问题[J]. 图书情报工作, 2014, 58(20): 13-22. (Wang Lianxi. Issues in Automatic Summarization Research [J]. Library and Information Service, 2014, 58(20): 13-22.)

[8] Nakov P I, Schwartz A S, Hearst M A. Citances: Citation Sentences for Semantic Analysis of Bioscience Text [C]. In: Proceedings of the SIGIR'04 Workshop on Search and Discovery in Bioinformatics, 2004: 81-88.

[9] Nanba H, Kando N, Okumura M. Classification of Research Papers Using Citation Links and Citation Types: Towards Automatic Review Article Generation [J]. Advances in Classification Research Online, 2000, 11(1): 117-134.

[10] Nanba H, Okumura M. Towards Multi-paper Summarization Using Reference Information [C]. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence, 1999: 926-931.

[11] 刘洋, 崔雷. 引文上下文在文献内容分析中的信息价值研究[J]. 图书情报工作, 2014, 58(6): 101-104. (Liu Yang, Cui Lei. The Information Value of Citation Context in Document Content Analysis [J]. Library and Information Service, 2014, 58(6): 101-104.)

[12] Qazvinian V, Radev D R. Identifying Non-explicit Citing Sentences for Citation-based Summarization [C]. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 2010: 555-564.

[13] Athar A, Teufel S. Detection of Implicit Citations for Sentiment Detection [C]. In: Proceedings of the Workshop on Detecting Structure in Scholarly Discourse, 2012: 18-26.

[14] 白光祖, 何远标, 马建霞, 等. 利用小样本量机器学习实现学术文摘结构的自动识别[J]. 现代图书情报技术, 2014(7-8): 34-40. (Bai Guangzu, He Yuanbiao, Ma Jianxia, et al. Application of Machine Learning with Limited Corpus to Identify Structure of Scientific Abstracts Automatically [J]. New Technology of Library and Information Service, 2014(7-8): 34-40.)

[15] Teufel S. Argumentative Zoning: Information Extraction from Scientific Text [D]. Edinburgh: University of Edinburgh School of Cognitive Science, 2000.

[16] Teufel S, Moens M. Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status [J]. Computational Linguistics, 2002, 28(4): 409-445.

[17] Liakata M, Dobnik S, Saha S, et al. A Discourse-Driven Content Model for Summarising Scientific Articles Evaluated in a Complex Question Answering Task [C]. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, USA. 2013: 747-757.

[18] Guo Y, Korhonen A, Liakata M, et al. Identifying the Information Structure of Scientific Abstracts: An Investigation of Three Different Schemes [C]. In: Proceedings of the 2010 Workshop on Biomedical Natural Language Processing (ACL 2010), 2010: 99-107.

[19] Mizuta Y, Korhonen A, Mullen T, et al. Zone Analysis in Biology Articles as a Basis for Information Extraction [J]. International Journal of Medical Informatics, 2006, 75(6): 468-487.

[20] Teufel S. Argumentative Zoning for Improved Citation Indexing [A]. //Computing Attitude and Affect in Text: Theory and Applications [M]. Netherlands: Springer, 2006: 159-169.

[21] Ehrler F, Geissbühler A, Jimeno A, et al. Data-poor Categorization and Passage Retrieval for Gene Ontology Annotation in Swiss-Prot [J]. BMC Bioinformatics, 2005, 6(S1): S23.

[22] Hirohata K, Okazaki N, Ananiadou S, et al. Identifying Sections in Scientific Abstracts Using Conditional Random Fields [C]. In: Proceedings of the International Joint Conference on Natural Language Processing, 2008: 381-388.

[23] Teufel S, Siddharthan A, Batchelor C. Towards Discipline-independent Argumentative Zoning: Evidence from Chemistry and Computational Linguistics [C]. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3, 2009: 1493-1502.

[24] Liakata M, Teufel S, Siddharthan A, et al. Corpora for the Conceptualisation and Zoning of Scientific Papers [C]. In: Proceedings of the International Conference on Language Resources and Evaluation, 2010: 2054-2061.

[25] Contractor D, Guo Y, Korhonen A. Using Argumentative Zones for Extractive Summarization of Scientific Articles [C]. In: Proceedings of the International Conference on Computational Linguistics, 2012: 663-678.

[26] Abu-Jbara A, Radev D. Coherent Citation-based Summarization of Scientific Papers [C]. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-

Volume 1, 2011: 500-509.
[27] Carbonell J, Goldstein J. The Use of MMR, Diversity-based Reranking for Reordering Documents and Producing Summaries [C]. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 1998: 335-336.
[28] Qazvinian V, Radev D R, Özgür A. Citation Summarization Through Keyphrase Extraction [C]. In: Proceedings of the 23rd International Conference on Computational Linguistics, 2010: 895-903.
[29] Mollá D, Jones C, Sarker A. Impact of Citing Papers for Summarisation of Clinical Documents[C]. In: Proceedings of the Australasian Language Technology Association Workshop, 2014: 79.
[30] Jaidka K, Chandrasekaran M K, Jha R, et al. The Computational Linguistics Summarization Pilot Task [C]. In: Proceedings of Text Analysis Conference, 2014.
[31] Radev D, Allison T, Blair-Goldensohn S, et al. MEAD-A Platform for Multidocument Multilingual Text Summarization [C]. In: Proceedings of Conference on Language Resources and Evaluation, 2004: 699-702.
[32] Chen J, Zhuge H. Summarization of Scientific Documents by Detecting Common Facts in Citations [J]. Future Generation Computer Systems, 2014, 32: 246-252.
[33] Galgani F, Compton P, Hoffmann A. Summarization Based on Bi-directional Citation Analysis [J]. Information Processing & Management, 2015, 51(1): 1-24.
[34] Erkan G, Radev D R. LexRank: Graph-based Lexical Centrality as Salience in Text Summarization[J]. Journal of Artificial Intelligence Research, 2004, 22: 457-479.
[35] Shi L, Tong H, Tang J, et al. VEGAS: Visual influEnce GrAph Summarization on Citation Networks [J]. IEEE Transactions on Knowledge and Data Engineering, 2015, 27(12): 3417-3431.
[36] Christensen J, Mausam S S, Soderland S, et al. Towards Coherent Multi-Document Summarization [C]. In: Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2013: 1163-1173.
[37] Barzilay R, Lapata M. Modeling Local Coherence: An Entity-based Approach [J]. Computational Linguistics, 2008, 34(1): 1-34.
[38] Lin C-Y. Rouge: A Package for Automatic Evaluation of Summaries [C]. In: Proceedings of the Workshop on Text Summarization Branches out. 2004.
[39] Nenkova A, Passonneau R. Evaluating Content Selection in Summarization: The Pyramid Method [C]. In: Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2004: 145-152.

作者贡献声明:

刘天祎：提出研究课题，文献阅读、整理，起草论文；
步一：文献阅读整理，起草论文；
赵丹群，黄文彬：论文最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据见期刊网络版 http://www.infotech.ac.cn。
[1] 刘天祎，步一，赵丹群，黄文彬．原始数据–检索结果.xlsx.
[2] 刘天祎，步一，赵丹群，黄文彬．原始数据–检索结果中相关文献.xlsx.

收稿日期: 2015-10-21
收修改稿日期: 2016-01-03

Review of Citation-based Automatic Summarization Studies

Liu Tianyi Bu Yi Zhao Danqun Huang Wenbin
(Department of Information Management, Peking University, Beijing 100871, China)

Abstract: [**Objective**] This paper is an in-depth review of popular research methodologies adopted by the Citation-Based Summarization (CBS) studies. [**Coverage**] We retrieved scholarly papers on CBS published since 2007, as well as earlier research on automatic summarization and citation analysis. [**Methods**] We thoroughly discussed the basic concepts and natural language processing technology in the field of CBS. [**Results**] Citances plays more important roles in automatic summarization applications than randomly selected sentences from scientific works. [**Limitations**] We did not compare the current achievements with possible results under the ideal circumstances. [**Conclusions**] CBS technology expands the scope of traditional informetrics and automatic summarization studies. It also offers suggestion to improve the existing evaluation methods of automatic summarization services. CBS calls for the expansion of citation windows and new experimental corpus. We have addressed these issues and explored new perspectives for the CBS research.

Keywords: Automatic summarization Citation-based summarization Citance Natural Language Processing